Section 2.2

Page 53, #2. An advantage of a stem-and-leaf plot is that it contains the original data points. A disadvantage is that it is harder to read than a histogram, and less flexible; for instance, widely dispersed data are very difficult to summarize with a stem-and-leaf plot.

Page 53, #8. max = 16.7, min = 12.9

Page 55, #24

If there is a relationship between hourly wage and hours worked, it is not readily apparent in this data set.

Section 2.3

Page 64, #4. False. Insight, p.59: “The mode is the only measure of central tendency that can be used to describe data at the nominal level of measurement.”

Page 64, #10. Skewed left or negatively skewed—the tail points to the left.

Page 65, #18. mean = 184.6, median = 182.5, mode does not exist. I would probably use the mean; although there are outliers, when it comes to blood cholesterol or other biomedical measurements, outliers are important! They tend to indicate sick people, and hence we want our measure of central tendency to be sensitive to them. But reasonable people can and do disagree on this issue.

Page 66, #32. $32,640

Page 67, #38. $1406/92 = 15.28
These data are right-skewed or positively skewed; the tail points to the right.

Section 2.4

Page 80, #10. The standard deviation is the square root of the variance. Neither can be negative—the variance is a sum of squares, and the standard deviation is defined to be the positive square root of the variance. The only way to get a standard deviation of zero is with homogeneous data, so the data set 7, 7, 7, 7, 7 would have \( n = 5, \bar{x} = 7, \) and \( s = 0. \)

Page 81, #16. Player B is more consistent because he has a smaller standard deviation of strokes, 1.2 vs. 2.3.

Page 82, #20. For public elementary school teacher salaries, range = 5.1, variance = 2.95, standard deviation = 1.72. For private school teachers, range = 5.2, variance = 1.99, and standard deviation = 1.41. Public school salaries are higher and slightly more variable—although we usually expect larger numbers to have larger variances. (We will learn about the coefficient of variation soon enough.)

Page 82, #26. About 95% of the data lie between $500 and $1900.

Page 83, #30. At least 75% of the data lie between 48.07 and 56.67.

Page 84, #38. The sample mean is approximately 43.7 and the sample standard deviation is approximately 23.7.

Page 85, #40. For male students the sample standard deviation is 127.4; for female students it is 185.9.

Section 2.5
Page 93, #8. False. The five numbers you need to graph a box-and-whisker plot are the minimum, the maximum, Q1, Q3, and the median.

Page 93, #12. min = 25, max = 85, Q1 = 50, Q2 (the median) = 65, Q3 = 70, IQR = 20.

Page 95, #26. The student has scored exactly the mean on both tests, so she has performed equally well on each test.

Page 95, #28a. 1st fly: \( z = 0.25 \), which corresponds to approximately the 60th percentile of the distribution. 2nd fly: \( z = -0.75 \), which corresponds to approximately the 23rd percentile. 3rd fly: \( z = 2.25 \), which corresponds to approximately the 99th percentile. I would call this last life span unusually long.

Extra Credit

1. \( xbar = 2271 \)

2. \( s = 653 \)

4. 

This distribution looks roughly bell-shaped, although somewhat truncated on the lower end.

5. 74% of the distribution lies within one standard deviation of the mean, and 98% lies within two standard deviations of the mean. This agrees fairly well with the Empirical Rule, although this distribution appears to have smaller tails than a bell-shaped or normal distribution.

6. The estimate of the mean using grouped data is 2317, which is close to the true sample mean of 2271.

7. The estimate of the standard deviation using grouped data is 642, which is close to the true sample standard deviation of 653.

8. The formulas for grouped data work well as long as there are a reasonable number of classes. Here we have seven classes. But imagine the problems we might have if we were only given two or three classes. In general we are better off working with the original data and using grouped data only as a last resort.